

An assistant for group formation in CSCL based on constraint satisfaction

Juan Bais, Jorge Giletto, José Balmaceda, Silvia Schiaffino and J. Andrés Díaz
Pace

ISISTAN (CONICET - UNCPBA), Campus Universitario, Paraje Arroyo Seco,
Tandil, 7000, Argentina
{jose.balmaceda, andres.diazpace, silvia.schiaffino}@isistan.unicen.edu.ar
{bais.juan, jsgiletto}@gmail.com

Abstract. Group formation is a key aspect in computer-supported collaborative learning, since different characteristics of students might influence the group performance. In this article, we present an assistant that models group formation as a weighted constraint satisfaction problem (WCSP), and considers three students' features, namely: psychological styles, team roles and social networks. Our WCSP formulation is able to combine constraints and preferences for individuals and groups. This assistant can aid teachers to form groups considering factors such as team role balance and distribution of psychological styles. We report on a pilot study to evaluate the proposal in different scenarios.

1 Introduction

Collaborative learning (CL) describes a situation in which certain types of interactions among people that promote their learning are expected, although not guaranteed [6]. The use of computational tools in the CL area has originated new teaching and learning scenarios, as well as new research opportunities. Computer-Supported Collaborative Learning (CSCL) aims at facilitating collaboration and communication among students with new technologies. In CSCL environments, students become independent of the time and space variables. That is, students can work collaboratively while situated in distant locations and even at different times.

When learning through CSCL tools, it is quite usual to work in groups. A group is a dynamic set of students that work together, discussing some topic, to eventually achieve some predefined goal. Each student of a group is responsible for his/her actions, but they work together on the same problem or exercise accepting the abilities and contributions of the other members. Having adequate groups/teams allows for a good interaction among the members and is fundamental in order to obtain appropriate learning results. Thus, group formation becomes a fundamental issue in CSCL.

To form groups, students can be either allocated to groups randomly, self-select each other, or be appointed to a group by the teacher based on some criteria related to the collaboration goals. These criteria are usually expressed

as a set of conditions, typically referred to as constraints, such as restricting the groups to be mixed in gender or skills [16]. For the teacher, forming groups manually can be both difficult and time consuming. For this reason, researchers have been investigating several techniques for automating this process through the use of computer-supported group formation (CSGF). Similar to manual group formation, the challenges of CSGF lie in modeling the students' features, the teacher's constraints; and negotiating the allocation of students to groups to satisfy these constraints.

In this work we propose an intelligent assistant for group formation in CSCL, which considers three different students' features that affect team performance: psychological styles, team roles, and social relationships. The assistant models the group formation problem as a weighted constraint satisfaction problem (WCSP). Psychological styles, as the model proposed by Myers and Briggs [14], can be considered as a way to model people's personality. Personality influences how a student works in a group. Team roles are a group of behavioral patterns expected and attributed to someone that occupies a certain position in a social unit. Several models and theories have emerged that study how the different roles contribute to the group work, and propose the different roles that people can take in a work group. In this work, we adopt the model proposed by Mumma [12]. Finally, it is known that students prefer to work with other students they already know or they have previously worked with. This information can be captured by the underlying social network of students.

The rest of the article is organized as follows. Section 2 presents an overview of our approach for automated group formation. In Section 3 we describe the main concepts related to Mumma's team roles theory. Similarly, in Section 4 we present a description of Myer-Briggs psychological styles. Then, in Section 5 we explain the formulation of group formation as a constraint satisfaction problem and we solve it. In Section 6 we present a case study, in which we used our approach to form groups in a CSCL context. In Section 7 we describe related work. Finally, in Section 8 we present the conclusions and future work.

2 Proposed approach

In order to guarantee that a team can achieve a good performance, as shown by previous works [10,17,20], it is necessary to have a diversity of psychological styles that assures that all aspects of a project will be addressed. A balance in the distribution of the different dimensions of Myer-Briggs psychological styles is necessary to achieve these goals [15]. On the other hand, according to Mumma [13], it is convenient to match the different team roles with the preferences of each of the team members. The author suggests that each team member has to adopt the role that is more convenient for him/her and for the team as a whole.

In this context, our approach considers the psychological characteristics and team role preferences of students as the basis to select team members, with the aim of forming teams that exhibit good performance and low level of conflicts. A general view of our proposal is shown in Figure 1. A user/student profile

consists of psychological styles and team roles. Also, information about a users' social network is used to consider the strength of relationships between him/her and other students. Conceptually, a social network is a structure composed of one or more graphs whose nodes represent actors or discrete social units, and edges represent relationships between them. Social network analysis enables us to obtain information about the relationships among individuals. Particularly, in our context, it is known that students tend to feel better working with people they already know or had already worked with before.

To model the desired features of teams according to the distribution of team roles and psychological styles, we propose a solution based on constraint satisfaction problems (CSP), particularly on weighted constraint satisfaction problems (WCSP). An example of constraint is that all team roles should be present in a team. User profiles, social networks and constraints are the inputs of our intelligent assistant. The outputs are the teams suggested for those inputs.

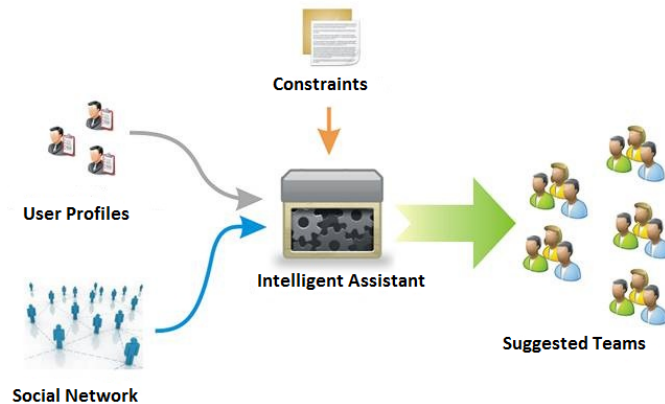


Fig. 1. Overview of our proposed approach

3 Team roles

The concept of role was defined as a group of behavioral patterns expected and attributed to someone that occupies a certain position in a social unit. A role refers to the way in which a team member interacts with others to facilitate group progress [3,5]. Several models and theories have emerged that study how the different roles contribute to the group work, and propose the different roles that people can take in a work group. Belbin [4] was the first researcher who proposed a team role theory. In the literature we can find several other team roles

models, apart from Belbin's, such as, Mumma [12], MTR-i¹, Insights², Thomas International³. Most of the models have developed tools that can determine team roles. These tools also provide reports that can help people to discover their current and potential abilities.

After a deep bibliographic review, we decided to use in our work the model proposed by Mumma [12]. This model is simple to apply since it provides a questionnaire that enables everyone to determine his/her preferred roles. Mumma's team role theory detects 8 different roles that can appear in a team. His theory is based on studies of Bales and Strodtbeck in 1953 that describe the different phases that a team suffers to solve a problem (as cited in [12]). Mumma observed that when groups move from one phase to another some roles become more important than others. Mumma defines 4 phases in the team work lifecycle. Each phase consists of two defined team roles, as described below.

- Phase 1- Initiation: It occurs when a task is defined. This task must be clearly stated along with its expected deliverable and allocated resources.
 - Role 1 - Leader: The leader inspires and motivates the team members.
 - Role 2 - Moderator: He/she matches the resources to the task at hand.
- Phase 2 - Ideation: It allows the team to identify alternative methods to perform a task such that needs can be fulfilled.
 - Role 3 - Creator: He/she identifies original ideas to approach a task along with alternatives.
 - Role 4 - Innovator: He/she identifies opportunities to use the various resources in the firm.
- Phase 3- Elaboration: This phase covers the elaboration of ideas invented from the ideation phase. The objective of this phase is to make the ideas work properly. Improper elaboration can cause conflicts with people, schedules, budgets and other resources.
 - Role 5 - Manager: He/she develops the plan to use resources and resolve conflicts.
 - Role 6 - Organizer: The organizer develops a plan to use time, money and resources such that the ideas created will work.
- Phase 4 - Completion: This phase covers the analysis of alternative methods, the decision of the plan of action and the execution of the task. Alternative methods to implement the task must be considered.
 - Role 7 - Evaluator: The evaluator makes judgments on situations, plans, results and alternatives.
 - Role 8 - Finisher: He/she follows plans and attends to the completion of the task.

¹ <http://www.rhassociates.com.au/mtri.htm>

² <http://www.insightinstitute.com/team-building-exercises.html>

³ <http://www.thomasinternational.net>

Mumma considers that sometimes teams fail at reaching their goals because people only carry out the tasks they like omitting some of the phases that are essential for the team progress. For this reason, it is important for a team to be composed of people having different team roles. Below, we describe the different team roles and their characteristics and responsibilities [12,13].

- *Leader*: he/she is the person who inspires and motivates the rest of the team members. The leader define, in general terms, the tasks that the team has to carry out, he/she ensures the achievement of these tasks and challenges the group to overcome the difficulties that might arise.
- *Moderator*: he/she identifies the capacities and abilities of each team member; he/she ensures the participation of all the members and assigns the resources for each task.
- *Creator*: he/she always generates ideas and alternatives to solve a certain problem. The creator finds innovative ways to overcome the difficulties that the team might encounter.
- *Innovator*: he/she is the person who identifies resources out of the team and finds opportunities to use them in tasks assigned to the team.
- *Manager*: he/she is the one that develops the plans to utilize human resources and solve the problems the team has. The manager keeps team members working cooperatively.
- *Organizer*: he/she develops the technical plans about the usage of time, money and physical resources, in order to materialize the ideas.
- *Evaluator*: he/she analyzes the situation, the plans, results and alternatives.
- *Finisher*: he/she follows the plans and checks that the tasks are carried out in time and correctly.

4 Pshychological styles

One possible way to determine the characteristics of a person's personality is by using the Myers-Briggs Type Indicator (MBTI). The MBTI assessment is a psychometric questionnaire designed to measure psychological preferences in how people perceive the world and make decisions⁴. These preferences were extrapolated from the typological theories proposed by Carl Jung in 1921, where Jung theorized that there are four principal psychological functions by which we experience the world: sensation, intuition, feeling, and thinking. One of these four functions is dominant most of the time.

The MBTI sorts some of these psychological differences into 4 opposite pairs, or dichotomies, which results into 16 possible psychological types. The 16 types are typically referred to by an abbreviation of four letters, the initial letters of each of their four type preferences (except in the case of intuition, which uses the abbreviation N to distinguish it from Introversion). For instance:

- ESTJ: extraversion (E), sensing (S), thinking (T), judgment (J)

⁴ http://www.personalitypathways.com/type_inventory.html

- INFP: introversion (I), intuition (N), feeling (F), perception (P)

Below, we describe the different dimensions of the MBTI [14,15].

- **Attitudes: extraversion/introversion (E/I)** Extraversion means "outward-turning" and introversion means "inward-turning". The preferences for extraversion and introversion are often called attitudes. People who prefer extraversion draw energy from action: they tend to act, then reflect, then act further. If they are inactive, their motivation tends to decline. To rebuild their energy, extraverts need breaks from time spent in reflection. Conversely, those who prefer introversion expend energy through action: they prefer to reflect, then act, then reflect again. To rebuild their energy, introverts need quiet time alone, away from activity.
- **Functions: sensing/intuition (S/N) and thinking/feeling (T/F)** Jung identified two pairs of psychological functions: two perceiving functions, sensing and intuition; and two judging functions, thinking and feeling. Sensing and intuition are the information-gathering (perceiving) functions. They describe how new information is understood and interpreted. Individuals who prefer sensing are more likely to trust information that is in the present, tangible, and concrete: that is, information that can be understood by the five senses. They prefer to look for details and facts. On the other hand, those who prefer intuition tend to trust information that is more abstract or theoretical, that can be associated with other information (either remembered or discovered by seeking a wider context or pattern). Thinking and feeling are the decision-making (judging) functions. The thinking and feeling functions are both used to make rational decisions, based on the data received from their information-gathering functions (sensing or intuition). Those who prefer thinking tend to decide things from a more detached standpoint, measuring the decision by what seems reasonable, logical, causal, consistent, and matching a given set of rules. Those who prefer feeling tend to come to decisions by associating or empathizing with the situation, and weighing the situation to achieve, on balance, the greatest harmony, consensus and fit, considering the needs of the people involved. Thinkers usually have trouble interacting with people who are inconsistent or illogical, and tend to give very direct feedback to others. They are concerned with the truth and view it as more important than being tactful.
- **Lifestyle: judging/perception (J/P)** Myers and Briggs added another dimension to Jung's typological model by identifying that people also have a preference for using either the judging function (thinking or feeling) or their perceiving function (sensing or intuition) when relating to the outside world (extraversion). Myers and Briggs held that types with a preference for judging show the world their preferred judging function (thinking or feeling). So TJ types tend to appear to the world as logical, and FJ types as empathetic. Those types who prefer perception show the world their preferred perceiving function (sensing or intuition). So SP types tend to appear to the world as concrete and NP types as abstract. According to this model perceptive types prefer to "keep decisions open".

5 Group formation as a constraint satisfaction problem

The assignment of persons to groups in such a way specific characteristics are fulfilled either at the individual level or at the group level can be naturally cast to a constraint satisfaction problem. A *constraint satisfaction problem* (CSP) is formally defined as a triple $\langle V, D, C \rangle$, where V is a set of variables, D is a domain of possible values for the variables, and C is a set of constraints over the values the variables can take. Every constraint is usually a boolean predicate $C_i(X) \rightarrow \{0, 1\}$ evaluated on a vector of variables X . When the domains of the variables are finite, CSP becomes a combinatorial problem that can be solved with several search techniques (e.g., backtracking, constraint propagation, local search). Thus, solving CSP consists of finding an assignment of all the variables to values in such a way all constraints are met. In our CSCL application, each slot of a group is a variable and its domain is the set of available people for the groups. The constraints derive from the models of Myer-Briggs and Mumma regarding psychological styles and team roles, respectively.

In real-world applications, there are often constraints of different kinds. Some constraints are mandatory and cannot be violated by feasible solutions, while other constraints affect the solution quality but do not need to be necessarily satisfied by solutions. The former are called *hard constraints* and the latter are called *soft constraints* (or preferences). An example of a hard constraint is that a person cannot belong to more than one group. An example of a soft constraint is that the members of a given group should be close in the underlying social network. When both hard and soft constraints are combined in the modeling of a problem, we have a variant of CSP referred to as *weighted constraint satisfaction* (WCSP). In WCSP, the evaluation of any constraint on a set of variables (bound to values) is associated to a weight, which often represents the cost (or penalty) of not satisfying that constraint. In our soft constraint above about closeness of team members, its cost can be a function of the size of the shortest paths between any two members. We can think of hard constraints as constraints whose cost is ∞ . Along this line, the goal of WCSP is not just satisfaction but also optimization. We have an objective function $F(A) = \sum_{j=1}^m C_j(A).cost_j(A)$ that sums the costs of evaluating every constraint C_i on a given variable assignment A . Thus, solving WCSP consists of finding an assignment for all variables in order to minimize such an objective function.

We used the Choco toolkit⁵ to implement the group formation and solve different problem instances. A problem instance is defined by the three parameters, namely: the list of available students (and their profiles), the number of groups K , and the size of each group N , as desired by the teacher. Conceptually, the solving process works in two stages. First, the solver searches for candidate solutions that meet all the hard constraints. Second, these solutions are assessed against the soft constraints. Finally, the solver outputs a ranking of p feasible solutions ordered by the total cost of the assignments in increasing order.

⁵ Choco homepage: <http://www.emn.fr/z-info/choco-solver/>

The main CSCL hard constraints that we modeled in Choco are the following:

- A user can belong to only one group.
- For each group, all the team roles of Mumma must be played by the members (one student can play more than one team role, according to Mumma's theory).
- Although a given role can be played by more than one team member, we limit this relationship to half the group size (e.g., in a group of size 6, up to 3 members can play the same role).
- For each dichotomy of Myers-Briggs, half of the members should prefer one function and the remaining members should prefer the other function, in order to favor heterogeneous groups.
- Permutations of persons and roles within the same group (i.e., symmetrical solutions) are not allowed.

The preferences of a given user for specific Mumma roles are stored in his/her profile. This profile also has a record of the psychological style applicable to the user. Based on this information, we defined the following soft constraints in Choco:

- The role preferences of any user must be balanced. According to [12], a user has balanced preferences if none role stands up over the other roles, or if the user does not dislike any role in particular. This relationship is quantified to a role score in the range [26-35]. Therefore, if some preferred roles are outside this range, a penalty is computed.
- Every user must have a psychological style compatible with the Mumma role she would like to realize, in order to perform efficiently. There is study ⁶ that maps the psychological styles of Myers-Briggs to corresponding Mumma roles for which a user would achieve a good fit.
- The distance in the social network for users assigned to the same group should be short. The shorter this distance, the lesser the penalty.

6 Case-study

We conducted a pilot study to evaluate the proposed approach. The dataset used for this study consisted of information about 44 students of a specialization course belonging to the 4th and 5th years of the Systems Engineering degree at UNCPBA, taught during 2011. These students completed Myers-Briggs and Mumma's questionnaires, to obtain their psychological styles and team roles, respectively. The information of social networks was also provided by the students, specifying relationships among them. Figure 2 shows a snapshot of our tool. On the left, the different students in the course are listed, and below, a group and its members is shown. On the top right, the characteristics of a selected student are displayed graphically.

⁶ http://www.belbinfrance.com/cms/cmsfiles/Later_versions_of_Belbin_Team_Roles.pdf

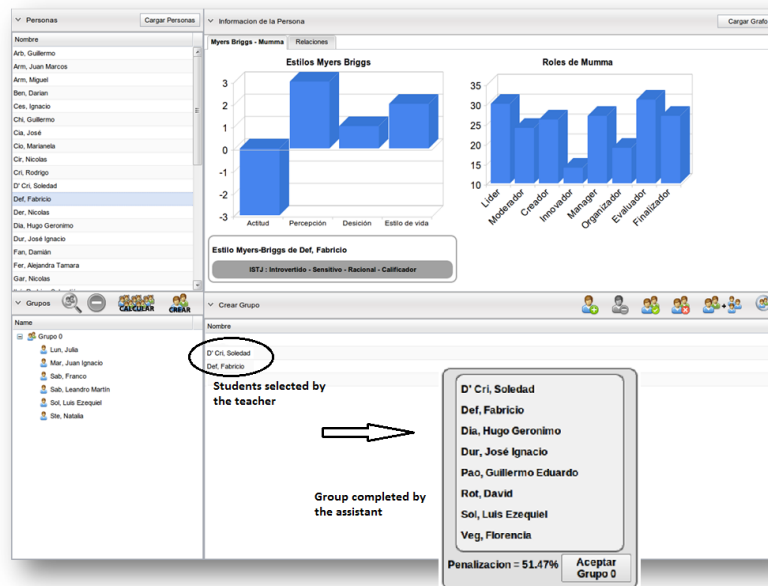


Fig. 2. Snapshot of the proposed assistant

The assistant can be used in different ways: i) to generate one group with a certain number of students; ii) to complete a group that already has some students; iii) to generate a certain number of groups having a fixed number of students. Figure 2 shows an example in which the teacher has selected two students for a group, and the assistant suggested the remaining students for that group. The penalty value is shown as a percentage of the total value of penalty a certain configuration can have. The lower the percentage, the better the group formation.

When generating for example, 6 groups with 7 students, the assistant shows the different solutions and the teacher might or might not accept each of the suggestions. When a suggested group is accepted, the students are eliminated from the database and the solver can generate new groups with the remaining students.

We also evaluated the execution times of the assistant, by considering the elapsed time between the moment in which a teacher requests a group formation and the moment when solutions are shown in the user interface. Figure 3 shows the results obtained when varying the number of groups and the number of team members. The test were executed on an Intel Dual Core 2.3 Mhz processor with 4 GB of memory. In general, the response times of the assistant do not exceed 11 seconds (for groups with 15 members). In average, the assistant suggests solutions in 2 to 6 seconds. This good performance is partially due to some built-in mechanisms of the Choco toolkit (e.g., support for global constraints, k-consistency techniques, etc.).

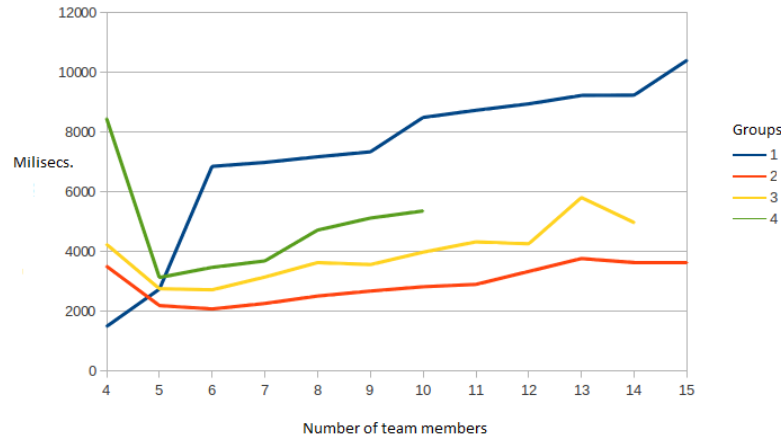


Fig. 3. Execution times of our tool

7 Related works

Some previous works have addressed the topic of group formation. OptAssign [9] is a web-based tool supporting both the workflow of collecting student data (preferences for some projects or topics) and the group formation. The latter is based on finding optimal solutions to suitable mathematical assignment problems, allowing for a number of constraints regarding size and structure of the groups. The evaluation results show advantages compared to manual procedures in terms of time savings for lecturers, and high fairness and correctness as perceived by students.

In [7] the authors formulate a group composition problem to model the formation of collaborative learning groups that satisfy the two grouping criteria. Moreover, this study is based on an approach called particle swarm optimization (PSO) to propose an enhanced PSO (EPSO) for composing well-structured collaborative learning groups.

In [16] a framework for learner group formation is proposed, based upon satisfying the constraints of the person forming the groups by reasoning over semantic data about the potential participants. The use of both Semantic Web technologies and Logic programming proved to increase the satisfaction of the constraints and overcome the orphans' problem. Zhamri Che Ani et. al. [2] present a method for group formation using a genetic algorithm, where the members for each group will be generated based on the students' programming skills in Java.

As regards other aspects considered for group formation, some other features are learning styles, argumentation abilities, students' interests, and other personality models. Felder-Silverman learning styles model is used in [1,8,11]. The first work, analyzes the effects of combining students with different learning styles on

group performance. Some rules based on the findings are also proposed. In [8], the authors propose the usage of Felder-Silverman model to form heterogeneous groups and automatically adapt proposed activities on a Web system. In [11] the influence of learning styles and argumentation capabilities of students on group performance is analyzed. With respect to personality [18,19], present models for group formation based on different personality tests. In [19], the authors provide some rules to group students considering their interests on a certain topic and Myers-Briggs psychological styles. In [18], some experiments are described where groups are formed combining students with different personalities.

Our work is novel in the sense that it combines psychological styles, team roles and social preferences of students. These factors have not been considered together thus far in previous works. In addition, modeling the problem as a WCSP is a new characteristic in group formation assistants. For instance, our formulation easily supports a ranking of assignment solutions for the teacher.

8 Conclusions and future work

In this article we have described an assistant for group formation based on constraint satisfaction, which considers three characteristics of students to form groups: psychological styles, team roles, and social networks. Modeling group formation as a WCSP is novel, as far as we are aware of, because it allows us to naturally integrate (and explore) constraints and preferences at the levels of individuals and groups. The proposed assistant can aid teachers to form groups in a CSCL context.

We conducted a pilot study to validate the proposal. We obtained groups for different scenarios and also evaluated execution times of the proposed approach. Although preliminary, these results are encouraging. As future lines of work, we will further evaluate our assistant by analyzing the performance of teams formed as suggested by our assistant, against the performance of teams formed randomly, or with other criteria. We will also compare our approach with other optimization techniques, such as genetic algorithms.

Acknowledgement

This work was partially supported by ANPCyT (Argentina) through PICT Project 2011 No. 0366 and PICT Project 2010 No. 2247, and also by CONICET (Argentina) through PIP Project No. 112-201101-00078.

References

1. E. Alfonseca, R. M. Carro, E. Martin, A. Ortigosa, and P. Paredes. The impact of learning styles on student grouping for collaborative learning: a case study. *User Modeling and User-Adapted Interaction*, 16:377–401, 2006.

2. Zhamri Che Ani, Azman Yasin, Mohd Zabidin Husin, and Zauridah Abdul Hamid. A method for group formation using genetic algorithm. *International Journal on Computer Science and Engineering*, 2(9):3060–3064, 2010.
3. A. Aritzeta, S. Swailes, and B. Senior. Belbin’s team role model: Development, validity and applications for team building. *Journal of Management Studies*, pages 96–118, 2007.
4. R. M. Belbin. *Management teams: Way they succeed or fail*. Oxford: Butterworth-Heinemann., 1981.
5. R. M. Belbin. *Team roles at work*. Oxford: Butterworth-Heinemann, 1993.
6. P. Jermann, A. Soller, and M. Mühlenbrock. From mirroring to guiding: A review of state of art technology for supporting collaborative learning. In *Proc. 1st European Conference on Computer Supported Collaborative Learning*, pages 324–331, 2001.
7. Yen-Ting Lin, Yueh-Min Huang, and Shu-Chen Cheng. An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers and Education*, 55:1483–1493, 2010.
8. E. Martín and P. Paredes. Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In *Proceedings ICWE Workshops*, pages 188–198, 2004.
9. D. Meyer. Optassign-a web-based tool for assigning students to groups. *Computers and Education*, 53:1104–1119, 2009.
10. R. Mitchell, S. Nicholas, and B Boyle. The role of openness to cognitive diversity and group processes in knowledge creation. *Small Group Research*, 40(5):535–554, 2009.
11. A. Monteserin, S. Schiaffino, P. Garcia, and A. Amandi. Análisis de la formación de grupos en aprendizaje colaborativo soportado por computadoras. In *Anais do XXI Simpósio Brasileiro de Informática na Educação, SBIE 2010*, 2010.
12. F. Mumma. *Team-Work & Team-Roles*. [1] King of Prussia: HRDQ, 1992.
13. F. Mumma. *Team-Work & Team-Roles: What makes your team tick?* Hrdq; 3rd edition, 2005.
14. I. Myers. *Manual: The Myers-Briggs Type Indicator*. Princeton, N.J.: Educational Testing Service, 1962.
15. I. Myers Briggs and P. B. Myers. *Gifts Differing: Understanding Personality Type*. Davies-Black Publishin, 1980.
16. A. Ounnas, H. C. Davis, and D. E Millard. A framework for semantic group formation in education. *Educational Technology & Society*, 12(4):43–55, 2009.
17. V. Pieterse, D. G. Kourie, and I. P Sonnekus. Software engineering team diversity and performance. In *Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries, SAICSIT ’06*, pages 180–186, 2006.
18. R. H. Rutherford. Using personality inventories to help form teams for software engineering class projects. In *In ITiCSE 01: Proceedings of the 6th Annual Conference on Innovation and Technology in Computer Science Education*, pages 73–76, 2001.
19. B. W. Speck. Fostering collaboration among students in problem-based learning. *New Direct. Teach. Learn*, 95:59–66, 2003.
20. M. F. Winter. *Developing a group model for student software engineering teams*. PhD thesis, 2004.